**Analysis on Suicide in the US from 2005 to 2016**

**Background**

Suicide in the United States of American has been trending on the various social media platforms and or in the news cycle in recent years. It is considered to be a major public health issue and amongst the top 10 leading cause of death in the US costing upwards of $60 billion annually. The number of suicides in the US has been increasing over the last several years and is currently at a 30 year high according to 2016 Center of Disease Control (CDC) data release. On average there are 123 suicides per day as reported and an estimate 1.4 million attempts. Relationship problems are one of the leading factors contributing to suicide.

Intrigued by this trend, the team of Jessica, Ellise, Yan and Escoffrey decided explore the topic of Suicide in America to gain a better understanding of the contributing factors and impact on the US population. The first question that came to mind is, what is suicide or how would you define suicide? The Merriam-Webster dictionary defines suicide as follows: noun “the act or instance of taking one’s life voluntarily and intentionally”; adjective “being or performing a deliberate act resulting in voluntary death of the person who does it” and as an intransitive verb “to commit suicide” or a transitive verb “to put oneself to death”.

Based on available information and our own assumptions the team set out to perform an analysis on the number of suicides death in the US over a period of time. In so doing, the team developed a hypothesis, a null hypothesis and several questions to address our hypothesis testing.

**Hypothesis**

If the number of individuals who died by suicide amongst millennials has increased since 2005, then the number of veterans who died by suicide within the same age group has increased.

**Null Hypotheses**

**Questions to address the Hypothesis**

To determine if we should reject the null hypotheses or fail to reject the null hypotheses in our analysis the following questions were raised:

* Amongst veterans is the suicide rate higher for females than males, and how does it compare to the general population (civilian population)?
* How does the age of veterans who have died by suicide vary in comparison to the general population (civilian population)?
* What methods are used most often by veterans who have died by suicide?
* Have significant events influenced suicide rates?

**Methods:**

To perform the analysis the following tools were used: Pandas Library for manipulation and analysis of the data; Matplotlib Library for plotting the results of the data analysis; Numpy Library for support of multi-dimensional arrays, matrices and the collection of high-level mathematical functions to operate on these arrays; Jupyter Notebook an integrated development environment (IDE) to write the codes and for data visualizations. Python program was used to perform (run) the overall analysis. Microsoft PowerPoint a presentation program to develop the presentation slides. Github a web based hosting service for version control using Git and source code management functionality.

The source data files Veterans Suicides by Sex, Suicides by Age, Suicides by Method,

Female and Male Suicides by Methods, General Population Suicide Methods, were retrieved on March 18, 2019 and downloaded as csv files form the CDC website: https://www.cdc.gov/violenceprevention/suicide/statistics.html and the VA Department website <https://www.mentalhealth.va.gov/mentalhealth/suicide_prevention/data.asp>? These files were read into Pandas using the Jupyter Notebook as separate dataframes. These dataframes were then manipulated by merging, dropping various columns and rows, converting data types and creating new data frames. The new dataframe files created were also saved as csv files to be used in the plotting graphs and charts to showcase the relationship between the respective variables being analyzed.

**Results and Observations:**

Based on the analysis performed in addressing the questions raised regarding suicides in the US between 2005 and 2016 the following results and observations were noted:

* Millennial total number of suicides: The results as shown by the bar chart “Veteran Suicide vs General Population”, comparing the number of suicides by veterans’ vs the general population for the Millennial (18-34) age group. Total suicides per veterans when from approximately 554 in 2005 to 895 in 2016 a 61.5% increase whereas the general population when from approximately 8,500 to 12,000 in 2016 a 41% increase.
* Millennial suicide percentage per 100,000: The results as shown by the bar chart “Percentage of Veteran Suicides vs General Population”, comparing the percentage of suicides by veterans’ vs the general population for the Millennial (18-34) age group. From 2005 to 2010 the percentage of veteran suicides per 100,000 was under 13% each year vs the general population of under 28%. For the period 2011 to 2016 veterans were between 13% and 17% vs the general population which was between 28% and 45%.
* Veteran suicide by sex: The number of veteran suicides was lower amongst female compared to males which was also true for the general population. From 2005 to 2016, more than 90% of veteran suicides were committed by males. As a note males made up just over 90% of the veterans population.
* Veteran suicide by age group: The results as shown by the line graph “Veteran Age Group Comparison”, the age group 18-34 had a dramatic increase from 25 per 100,000 to 45 per 100,000 suicides. This was and 80% increase over the period. Also, it was noted that this age group recorded the highest suicide rate from 2013 to 2016.
* General population suicide by age group: Suicide appears to be more prevalent in the “35-54” age group as compared to the other age groups in the general population. It was noted that all age groups showed a gradual increase in rates from 2006 to 2015. This can be seen on the line char “General Population Age Group Comparison”.
* Total suicides percent 2005 – 2016: The results as depicted by the line chart “2005 – 2016 Suicides by Percentage” showing the percentage of total suicides of the civilian vs the veteran population. While the civilian population has a higher suicide count in raw numbers, the risk of suicide was higher amongst veterans. Additionally, each age group had an increase in the number of suicides, the increase in veterans were 6% compared to 2% for the civilian population.
* Methods of used: As visualized in the respective stack bar charts, “Veteran Deaths by Method” and “Civilian Deaths by Method” both veterans and civilians favor to use firearms as their method of suicide. Firearm was the method of choice used by veterans in more than 60% of suicides compared to an average of less than 50% for civilians.
* Additional analysis: As depicted in the bar chart “Veteran Suicides by Region from 2005 to 2016” the number of suicides by veterans was greatest in the southern region.
* Significant events between 2005 and 2016. During the period the following significant events occurred in the US:
  + 2005 Hurricane Katrina (Natural Disaster)
  + 2007 Shooting on VA Tech Campus (Mass Shooting)
  + 2008 Great Recession – property and asset values plummeted , stock market crash, liquidity issues and mass layoffs (Economy)
  + 2009 Great Recession continued (Economy); Shooting at Fort Hood (Mass Shooting)
  + 2010 Unemployment rate rises to 9.6% (Economy)
  + 2011 Hurricane Irene (Natural Disaster)
  + 2012 Hurricane Sandy (Natural Disaster)
  + 2012 Sandy Hook Elementary School and Aurora Colorado and Okis University (Mass Shootings)
  + 2013 Boston Marathon Bombing (Domestic Terrorism)
  + 2014 Isla Vista Valley Killings (Mass Shooting)
  + 2015 Umpqua Community College and San Bernardino Attack (Mass Shootings)
  + 2016 Orlando Nightclub Shooting (Mass Shooting)

**Limitations**

In conducting our analysis several limitations were noted. The number of suicides each year may be under reported due to the associated stigma and the classification of death. Due to the lack of data, the rate of suicide amongst veterans by sex was unobtainable. Given the source of the data used we were not able to perform a meaning fully analysis by state or by ethnicity. The analysis was over a limited period of 11 years. No p-value test and or regression analyst was performed.

**Opportunities**

From the analysis performed, these opportunities were noted. Educate the entire population on suicides and overcoming the associated stigma. Track suicides across the entire US population to describe trends, circumstances, and populations at greatest risk. Have more trained professionals, making access to those in need more readily available and affordably. Train providers in adopting proven treatments for patients at risk of suicide. Perform more studies focusing on developing, implementing and evaluating suicide prevention strategies. Teach coping and problem-solving skills to help people manage challenges with relationships, jobs, health, or other concerns. Promote safe and supportive environments such as safely storing medications and firearms to reduce access among people at risk. Offer activities that bring people together so they feel connected and not alone. Encourage employers to promote employee health and well-being, support employees at risk, and have plans in place to respond to people showing warning signs.

**Conclusion**

Female Veterans are more likely to die by suicide than females in the general population. Millennials accounts for highest percentage change in suicide rates with the millennials veterans’ population recording the most significant increase over the period. The preferred method by both veterans and the general population is firearms. Although, males account form the vast majority of suicides each year, the number of suicides committed by female is increasing each year noting that the suicide rate of female veterans was nearly twice the suicide rate of females in the general population.

Based on the results of our analysis and the trends and limitations noted, we were not able to come to a conclusion to reject or accept the null. Further analysis need to be performed in order to such a conclusion.